Method for the analysis of gynaecological cell proliferative disorders

The present invention relates to a method for the detection, differentiation and prognosis of a gynaecological cell proliferative disorder, comprising the following steps: a) obtaining a cervicovaginal secretion specimen from an individual b) determining the methylation status of at least one or more CpG positions c) determining from said methylation status the presence, classification and/or prognosis of a gynaecological cell proliferative disorder in said individual. The present invention further relates to a kit for performing said method.

Field of the Invention

Endometrial cancer

Endometrial cancer is one of the most common genital cancers in women worldwide. The highest incidence rates are observed in Western Europe and North America. The well known risk factors for endometrial cancer include obesity, type 2 diabetes mellitus and hypertension. Additionally, anovulation and long term use of unopposed estrogens for hormone replacement therapy increase the risk for endometrial cancer. Genetic causes of endometrial cancer are uncommon, although there is an association with hereditary non-polyposis colon cancer syndrome, in which the individual risk rises to a cumulative incidence of 40% by age 70 years. In 2001, the American Cancer Society concluded that there was insufficient evidence to recommend screening for endometrial cancer for women at average risk or increased risk due to history of unopposed estrogen therapy, nulliparity, infertility or failure to ovulate, obesity, diabetes, or hypertension. Studies examining endometrial carcinoma screening methods for asymptomatic postmenopausal women have used ultrasound-determined endometrial thickness as an indication of risk. Transvaginal ultrasonography compared with endometrial biopsy for the detection of endometrial disease had a positive predictive value of only nine percent for detecting any abnormality, with 90 percent sensitivity and 48% specificity (Langer, R. D., Pierce, J. J., O'Hanlan, K. A., Johnson, S. R., Espeland, M. A., Trabal, J. F., Barnabei, V. M., Merino, M. J., and Scully, R. E. Transvaginal ultrasonography compared with endometrial biopsy for the detection of endometrial disease. Postmenopausal Estrogen/Progestin Interventions Trial. N. Engl. J. Med., *337:* 1792-1798, 1997.). There is a need for a sensitive and specific screening test for high-risk women. It has been shown that genetic abnormalities can be used to detect endometrial cancer. Changes in the status of DNA methylation are among the

most common molecular alterations in human neoplasias (Jones, P. A. DNA methylation errors and cancer. Cancer Res., *56:* 2463-2467, 1996.). It has been increasingly recognized over the past four to five years that the CpG islands of a large number of genes, which are unmethylated in normal tissue, are methylated to varying degrees in multiple types of human cancer (Jones, P. A. and Laird, P. W. Cancer epigenetics comes of age. Nat. Genet., *21:* 163-167, 1999). Aberrant methylation of CpG islands within the promoter regions of several genes such as *E-cadherin, adenomatous polyposis coli* (*APC*), *MLH1, p16, estrogen receptor, progesterone receptor* and *PTEN* (*MMAC1*) has been identified in endometrial cancer tissue. Up to now, no investigations have been undertaken to assess the methylation status of DNA obtained from cervicovaginal secretion from endometrial cancer patients.

Long-term tamoxifen users who are at increased risk for endometrial cancer have a worse prognosis for such cancers, which seems to be due to less favorable histology and higher stage. This indicates an urgent need for a simple, non-invasive means of early detection of endometrial cancer, especially in this subgroup of women.

Cervical cancer

Cancer of the uterine cervix is an important cause of death in women worldwide. Since the introduction of PAP smears in screening programs the incidence and mortality of cervical cancer have decreased dramatically. However, successful screening strongly depends on the coverage rate of the population and the sensitivity and specificity of the screening test. A meta-analysis of studies investigating the pap test for the detection of cervical cancer and its precursors revealed a sensitivity ranging from 30% to 87% and a specificity ranging from 86% to 100%. Converging evidence from epidemiological and molecular studies suggests that infection with genital human papillomavirus (HPV[3]) is causally linked to the development of cervical cancer. Therefore, testing for HPV DNA has been evaluated to improve cervical cancer screening. Numerous studies showed a high sensitivity for the HPV test in detecting cervical cancer and its precursors, whereas specificity was usually lower in comparison to cytology. To reduce the inconvenience and cost of repeated clinical visits, it has been proposed that women collect cervicovaginal specimens themselves for HPV DNA assay, hopefully increasing the coverage of screening programs.

Several studies investigated HPV DNA detection rates between self-collected and physician-collected samples with varying concordance between the two different collection methods.

In addition to HPV infection, it is clear that other factors are also involved in cervical carcinogenesis because the majority of patients with HPV-associated lesions do not progress to invasive cancer. Changes in the status of DNA methylation are among the most common molecular alterations in human neoplasias. Recently, an aberrant methylation pattern was found during the multistage pathogenesis of cervical cancer with a trend to increasing methylation with increasing pathological changes (Virmani, A. K., Muller, C., Rathi, A., Zoechbauer-Mueller, S., Mathis, M., and Gazdar, A. F. Aberrant methylation during cervical carcinogenesis. Clin. Cancer Res., 7: 584-589, 2001).

In a study by Kinney *et al.* Up to 60% of women diagnosed as having invasive cervical cancer were not screened despite health maintenance organization enrolment. To reduce the inconvenience and cost of repeated clinical visits it has been proposed that women collect cervicovaginal specimens themselves, thus hopefully increasing the coverage of screening programs. Previous studies investigating HPV DNA detection rates between self-collected and physician-collected samples described varying concordance between the two different collection methods. Recently, it was shown that clinician-directed swabs detect up to 28% more HPV-positive women in comparison to tampon-collected specimens. In light of the fact that numerous studies have revealed a nearly 100% sensitivity in detecting SIL and cervical cancer by HPV DNA testing of physician-collected samples, specimen collection by tampon seems not to be a feasible method for HPV DNA detection.

It has been proposed that in addition to HPV infection genetic or epigenetic alterations may be required to maintain a malignant phenotype. Changes in the status of DNA methylation are among the most common molecular alterations in human neoplasias. Recently it was suggested that aberrant methylation may play a role in cervical carcinogenesis (Virmani, A. K., Muller, C., Rathi, A., Zoechbauer-Mueller, S., Mathis, M., and Gazdar, A. F. Aberrant methylation during cervical carcinogenesis. Clin. Cancer Res., 7: 584-589, 2001).

Cervical cancer is the principal cause of death due to cancer in women. Five-year survival rate ranges from 15 to 80 percent, depending on the extent of the disease. Recently, several studies showed a significant reduction in the risk of relapse and death from cervical cancer, which was achieved by concurrent use of chemotherapy and radiotherapy. New predictive markers for relapse may increase survival rates by improving treatment of patients at high risk for relapse.

Several clinical and histopathological characteristics, namely tumor stage, lymph node metastasis and vascular invasion, have been shown to be prognostic factors for recurrent disease. However, new molecular and biochemical approaches for the recognition and treatment of high risk patients are needed to improve survival and avoid over-treatment of low-risk patients. The gene products of *CDH1* and *CDH13*, namely E-cadherin and H-cadherin, play a key role in cell-cell adhesion. Changes in cell-cell and cell-matrix adhesion accompany the transition from benign tumor to invasive, malignant cancer and the subsequent metastatic dissemination of tumor cells. Decrease or loss of E-cadherin expression is a common finding in many human epithelial cancers including cervical cancer. The cadherin-mediated cell adhesion system can be inactivated by several mechanisms. It has been reported that aberrant methylation of CpG islands in the E-cadherin (*CDH1*) as well as in the H-cadherin (*CDH13*) promotor or 5'-region may lead to decreased E-cadherin and H-cadherin expression. Numerous studies have demonstrated tumor-specific alterations in DNA recovered from plasma or serum of patients with various malignancies, a finding that has potential for molecular diagnosis and prognosis.

Abnormalities of cell adhesion molecule expression like E-cadherin (*CDH1*) and H-cadherin (*CDH13*) occur in various neoplastic diseases, and there is some evidence to suggest that these abnormalities are significant in the progression of certain tumor types including cervical cancer. Several mechanisms like tumor hypoxia and necrosis, stimulation of the epidermal growth factor receptor (EGFR) by EGF or TGF-α and mutations of the CDH1 gene have been proposed for cadherin downregulation. Recently, aberrant promoter methylation of *CDH1* and *CDH13* has been described to be one of the mechanisms causing loss of or decreased E-cadherin and H-cadherin expression. Decreased E-cadherin expression has been shown to be related to enhanced metastasizing activity or more aggressive malignant tumors.

Methylated DNA has been investigated as a possible screening marker for neoplastic disease in several body fluids (Muller, H. M. and Widschwendter, M. Methylated DNA as a possible screening marker for neoplastic disease in several body fluids. Expert Rev Mol Diagn, *3:* 443-458, 2003). However, up to now, no investigations have been undertaken to assess the methylation status of DNA obtained from cervicovaginal secretion for the assessment of patients with gynaecological cellproliferative disorders.

Description

Current methods of endometrial cancer diagnosis, namely transvaginal ultrasonography compared with endometrial biopsy for the detection of endometrial disease are estimated to have a positive predictive value of only nine percent for detecting any abnormality, with 90% sensitivity and 48% specificity. There is a need in the art for a sensitive and specific screening test for high-risk women.

The standard screening test for cervical cancer is the PAP smear. Successful screening strongly depends on the coverage rate of the population and the sensitivity and specificity of the screening test. A meta-analysis of studies investigating the pap test for the detection of cervical cancer and its precursors revealed a sensitivity ranging from 30% to 87% and a specificity ranging from 86% to 100%. However the success of screening programs aimed at detecting precancerous conditions (dysplasia) and treating them before they progress is often limited by socio-economic factors. It has been estimated that only about 5 % of women in developing countries have been screened for cervical dysplasia in the past 5 years, compared with some 40% to 50% of women in developed countries. In order to improve the successful screening of populations for cervical cancer, especially of the high-risk population with low socio-economic status, there is a need for a sensitive, specific, cost effective and self-administrable cervical cancer test.

Cervical cancer is in many cases a treatable disease, however for those patients who have poor prognosis a significant reduction in the risk of relapse and death can be achieved by concurrent use of chemotherapy and radiotherapy. However, current prognostic markers, mainly histological, only provide a limited indication of patient prognosis. Furthermore treatment of patients who are not at risk of relapse by adjuvant treatments can lead to unnecessary side effects. Therefore there exists a need in the art for improved and preferably self-administrable means of detection and prognosis of cervical cancer.

As used herein the term "prognosis" shall be taken to mean a prediction of the progression of the disease, it may be measured by reference to any suitable parameters including but not limited to survival. It is preferably used to help define patients with high and low risks of death that result from the inherent heterogeneity of a disease process.

As used herein the term "survival" shall be taken to include survival until mortality (wherein said mortality may be either irrespective of cause or gynaecological cell proliferative disorder related); recurrence free survival (wherein the term recurrence shall include both localized and distant recurrence); metastasis free survival and disease free survival (wherein the term disease shall include gynaecological cell proliferative disorders and diseases associated therewith). The length of said survival may be calculated by reference to a user defined start point (e.g. time of diagnosis or start of treatment) and end point (e.g. death, recurrence or metastasis).

As used herein the term "cervicovaginal secretion" shall be taken to mean any substance discharged from any cell, gland, or organ into the cervical or vaginal areas.

The method according to the invention provides novel sensitive and specific means for the detection, differentiation and prognosis of gynaecological cell proliferative disorders by analysis of cervicovaginal secretions. The method thereby enables specimen collection by the patient without the need for the aid of a physician. In a further aspect the invention provides methods for the analysis of said specimen for genomic methylation features associated with a variety of gynaecological cell proliferative disorders, in particular those of the endometrium and cervix. In a further aspect the invention provides novel nucleic acids for the analysis of said cervicovaginal secretion specimen that enable the detection, differentiation and prognosis of gynaecological cell proliferative disorders. The invention further provides kits for the analysis of cervicovaginal secretions for the analysis of methylation features that enable the detection, differentiation and prognosis of gynaecological cell proliferative disorders.

The method according to the invention comprises the following steps:

a) obtaining a cervicovaginal secretion specimen from an individual,
b) determining the methylation status of at least one or more CpG positions,
c) determining from said methylation status the presence, classification and/or prognosis of a gynaecological cell proliferative disorder in said individual.

The cervicovaginal secretion maybe obtained by any means standard in the art, including but not limited to gynaecological swab, aspiration and cervicovaginal lavage. However in the most preferred embodiment of the method said specimen is collected by means of a tampon inserted into the individual's vaginal passage, it is preferred that the tampon is then stored in a

buffer or other suitable solution. It is a particularly preferred embodiment of the invention that this step of the method is carried out by said individual, without the assistance of a physician or other healthcare professional. The solute containing DNA from the cervicovaginal may then be used in the further steps of the method.

The utility of self administered cervicovaginal specimen collection has been well established for testing of HPV (Human papilloma virus), an indicator of cervical cancer. Studies in the United States (Harper DM, Hildesheim A, Cobb JL, Greenberg M, Vaught J, Lorincz AT. Collection devices for human papillomavirus. J Fam Pract. 1999 Jul;48(7):531-5.) showed an 80% diagnosis concordance rate between specimens collected by the physician and those collected by the patient using vaginal tampons. However, the use of this technique for analysis of the diseased tissues themselves have so far, not been published.

The cervicovaginal specimen, or solution thereof is then analysed by means of a methylation assay in order to detect aberrant methylation patterns associated with the development of gynaecological cell proliferative disorders, this is enabled by analysing one or more CpG positions the aberrant methylation of which is a feature of a gynaecological cell proliferative disorders or an indicator of prognosis thereof. In the final step of the method the presence, classification and/or prognosis of a gynaecological cell proliferative disorder is determined from said determined methylation status of one or more CpG positions.

In a preferred embodiment of the method the gynaecological cell proliferative disorder is selected from the group consisting no dysplasia or low grade squamous intraepithelial lesions, high-grade squamous intraepithelial lesions, cervical cancer, endometrial cancer and grade 1 to 3 cervical intraepithelial neoplasia.

In one embodiment of the method said CpG positions are selected from one or more genes taken from the group consisting of CDH1, CDH13, RASSF1A, hMLH1, HSPA2, SOCS1, SOCS2, GSTP1, DAPK, TIMP3, hTERT, SFRP2, SFRP4, SFRP5 and CCND2 and/or promoters, introns, first exons, regulatory elements and/or enhancers thereof. It is also a further embodiment of the invention that the sequence of said genes are selected from the group consisting of SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 according to Table 7.

In one embodiment of the method endometrial cancer is detected or differentiated from other endometrial cell proliferative disorders by analysis of one or more CpG positions of the genes *SFRP2, SFRP4, SFRP5, CCND2, RASSF1A, hMLH1, CDH13, HSPA2* and *SOCS2*, in a further embodiment of said method the sequences of said genes are selected from SEQ ID Nos: 2, 3, 4, 5, 7, 64, 65, 66 and 67 according to Table 7.

In one embodiment of the method cervical cancer is detected or differentiated from other cervical cell proliferative disorders by analysis of one or more CpG positions of the genes *SFRP2, SFRP4, SFRP5, CCND2, SOCS1, CDH1, TIMP3, GSTP1, DAPK, hTERT, CDH13, HSPA2, MLH1, RASSF1A* and *SOCS2*, in a further embodiment of said method the sequences of said genes are selected from SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 according to Table 7.

In a further embodiment of the method the prognosis of a patient with cervical cancer is determined by analysis of one or both of the genes *CDH1* and *CDH13*, in a further embodiment of said method the sequences of said genes are selected from SEQ ID NO:1 and 2 according to Table 7. Hypermethylation of the genes are correlated with a worse prognosis of life expectancy of the patient.

Once a cervicovaginal secretion specimen is obtained from the patient according to step a) of the method genomic DNA is isolated. This may be by any means standard in the art, including the use of commercially available kits. Briefly, wherein the DNA of interest is encapsulated by a cellular membrane the biological sample must be disrupted and lysed by enzymatic, chemical or mechanical means. The DNA solution may then be cleared of proteins and other contaminants e.g. by digestion with proteinase K. The genomic DNA is then recovered from the solution. This may be carried out by means of a variety of methods including salting out, organic extraction or binding of the DNA to a solid phase support. The choice of method will be affected by several factors including time, expense and required quantity of DNA.

The isolated genomic DNA is then analysed for aberrant methylation of CpG dinucleotides. This may be by any means standard in the art including the use of methylation sensitive restriction enzymes. However, it is required that a method capable of high sensitivity be used as there is only likely to be a minimal amount of DNA from the diseased tissue present in the sample. Therefore it is particularly preferred that step b) of the method is carried out by treat-

ing the genomic DNA, or a fragment thereof, with one or more reagents to convert 5-position unmethylated cytosine bases to uracil or to another base that is detectably dissimilar to cytosine in terms of hybridization properties. This treatment is preferably carried out by means of a bisulfite reagent (bisulfite, disulfite, hydrogen sulfite or combinations thereof) followed by alkaline hydrolysis. It is a further embodiment of the method that post-treatment the sequences to be analyzed are taken from the group consisting of SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83.

The treated genomic DNA, or the treated fragment thereof, are then preferably contacted with an amplification enzyme and at least two primers comprising, in each case a contiguous sequence at least 9 nucleotides in length that is complementary to, or hybridizes under moderately stringent or stringent conditions to a target nucleic acid. It is a further embodiment of the method that post-treatment the sequences to be analyzed are taken from the group consisting of SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83. In the final step of the method the methylation state of at least one CpG dinucleotide sequence, or an average, or a value reflecting an average methylation state of a plurality of CpG dinucleotide sequences is determined based on a presence or absence of, or on a property of said amplificate whereby at least one of detecting a gynecological cell proliferative disorder, or distinguishing between gynecological cell proliferative disorders is, or providing a prognosis at least in part, afforded.

This may be carried out by any means standard in the art that enables the detection of small amounts of methylated DNA against a high background of non-methylated DNA. Particularly preferred are the MSP, HeavyMethyl (blocking oligonucleotides) and RealTime assays (including, but not limited to the Ligthcycler and MethyLight assays) and all possible combinations thereof.

The term "MSP" (Methylation-specific PCR) refers to the art-recognized methylation assay described by Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996. In MSP applications, the use of methylation status specific primers for the amplification of bisulphite-treated DNA allows for distinguishing between methylated and unmethylated nucleic acids. MSP primer pairs contain at least one primer which hybridizes to a bisulphite-treated CpG dinucleotide of a pre-specified methylation state. Therefore, the sequence of said primers comprises at least one CpG, TpG or CpA dinucleotide. MSP primers specific for non-methylated

DNA contain a 'T' at the 3' position of the C-position in the CpG dinucleotide. Detection of the amplificate allows for the determination of the presence of a methylated nucleic acid. The use of MSP thereby allows for the detection of a nucleic acid of a pre-specified methylation state to be amplified against a background of alternatively methylated nucleic acids (*see* figure 8 herein and the accompanying description).

In the HeavyMethyl® technique, polymerase amplification is aided by the use of blocking oligonucleotides, this technique may be used to block amplification of target nucleic acids of one methylation status thereby increasing the relative proportion of the amplificate nucleic acids of the other methylation status. The primers may be methylation specific ('MSP') or specific to non-CpG treated target nucleic acids. The methylation status of the bisulphite-treated CpG dinucleotides is determined by means of oligonucleotide blocking probes that are not displaced by the action of the polymerase, and thus block amplification of the sequence (*see* Figure 9). Non-displacement of the blocking oligonucleotides may be achieved by use of a polymerase that has no 5'-3' exonuclease activity, by use of peptide nucleic acid oligomers or by use of suitably modified oligonucleotides (e.g. a DNA oligomer lacking a free 3'-hydroxyl group)

Figure 9 shows polymerase-mediated amplification analysis of bisulfite-treated DNA ("3") corresponding to a CpG-rich genomic sequence by means of the HeavyMethyl® technique. Amplification of the treated DNA ("3") is precluded if the blocking oligonucleotide ("5") anneals to the treated DNA as shown for the example case "B." The arrows ("1") represent primers, and dark circular marker positions ("2") on the bisulfite-treated nucleic acid strand ("3") represent methylated bisulfite-converted CpG positions, whereas white circular marker positions ("4") represent unmethylated bisulfite-converted positions. The blocking (blocker) oligonucleotides are represented by dark bars ("5"). In the example case "A," all subject genomic CpG positions were co-methylated, and both forward and reverse primers anneal to provide for unimpeded amplification of the corresponding treated nucleic acid ("3"). In the second example case "B," none of the subject genomic CpG positions were methylated, both forward and reverse primers anneal to the treated DNA sequence ("3") but are unable to amplify the sequence, because the synthesis of the complementary strand is blocked by the blocking oligonucleotide ("5") that anneals to a complementary position comprising unmethylated CpG sequences in the subject genomic DNA.

'Real time' based methods employ the use of a detection probe that hybridizes to a treated target nucleic acids during PCR amplification thereby enabling the detection of the amplificate nucleic acid. The detection probes may be designed to hybridize to methylation specific target sequences by comprising one or more CpG, TpG or CpA dinucleotides. Suitable real time PCR based methods, include the art-recognized fluorescence-based real-time PCR technique MethyLight™ (Eads et al., *Cancer Res.* 59:2302-2306, 1999; U.S. Patent No. 6,331,393 to Laird et al.; *and see* Heid et al., *Genome Res.* 6:986–994, 1996) and the TaqMan assay. A particularly preferred embodiment comprises use of fluorescence-based Real Time Quantitative PCR (Heid et al., *Genome Res.* 6:986-994, 1996) employing a dual-labeled fluorescent oligonucleotide probe (TaqMan™ PCR, using an ABI Prism 7700 Sequence Detection System, Perkin Elmer Applied Biosystems, Foster City, California). The TaqMan™ PCR reaction employs the use of a nonextendible interrogating oligonucleotide, called a TaqMan™ probe, which is designed to hybridize to a CpG-rich sequence located between the forward and reverse amplification primers. The TaqMan™ probe further comprises a fluorescent "reporter moiety" and a "quencher moiety" covalently bound to linker moieties (*e.g.*, phosphoramidites) attached to the nucleotides of the TaqMan™ oligonucleotide. For analysis of methylation within nucleic acids subsequent to bisulfite treatment, the probe is preferably methylation specific, as described in U.S. 6,331,393, (hereby incorporated by reference) also known as the MethylLight® assay. Variations on the TaqMan™ detection methodology that are also suitable for use with the described invention include the use of dual probe technology (Lightcycler™) or fluorescent amplification primers (Sunrise™ technology).

The disclosed invention provides treated nucleic acids, derived from genomic SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67, wherein the treatment is suitable to convert at least one unmethylated cytosine base of the genomic DNA sequence to uracil or another base that is detectably dissimilar to cytosine in terms of hybridisation. The genomic sequences in question may comprise one, or more, consecutive or random methylated CpG positions. Said treatment preferably comprises use of a reagent selected from the group consisting of bisulfite, hydrogen sulfite, disulfite, and combinations thereof. In a preferred embodiment of the invention, the objective comprises analysis of a non-naturally occurring modified nucleic acid comprising a sequence of at least 16 contiguous nucleotide bases in length of a sequence selected from the group consisting of SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83, wherein said sequence comprises at least one CpG, TpA or CpA dinucleotide and sequences complementary thereto. The sequences of SEQ ID

NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83 provide non-naturally occurring modified versions of the nucleic acid according to SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 and SEQ ID NO: 64 to SEQ ID NO: 67, wherein the modification of each genomic sequence results in the synthesis of a nucleic acid having a sequence that is unique and distinct from said genomic sequence as follows. For each sense strand genomic DNA, e.g., SEQ ID NO: 1, four converted versions are disclosed. A first version wherein "C" is converted to "T," but "CpG" remains "CpG" (i.e., corresponds to case where, for the genomic sequence, all "C" residues of CpG dinucleotide sequences are methylated and are thus not converted); a second version discloses the complement of the disclosed genomic DNA sequence (i.e. antisense strand), wherein "C" is converted to "T," but "CpG" remains "CpG" (i.e., corresponds to case where, for all "C" residues of CpG dinucleotide sequences are methylated and are thus not converted). The 'upmethylated' converted sequences of SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 correspond to SEQ ID NO: 12 to SEQ ID NO: 33 and SEQ ID NO: 68 to SEQ ID NO: 75. A third chemically converted version of each genomic sequences is provided, wherein "C" is converted to "T" for all "C" residues, including those of "CpG" dinucleotide sequences (i.e., corresponds to case where, for the genomic sequences, all "C" residues of CpG dinucleotide sequences are unmethylated); a final chemically converted version of each sequence, discloses the complement of the disclosed genomic DNA sequence (i.e. antisense strand), wherein "C" is converted to "T" for all "C" residues, including those of "CpG" dinucleotide sequences (i.e., corresponds to case where, for the complement (antisense strand) of each genomic sequence, all "C" residues of CpG dinucleotide sequences are unmethylated).The 'downmethylated' converted sequences of SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 correspond to SEQ ID NO: 34 to SEQ ID NO: 55 and SEQ ID NO: 76 to SEQ ID NO: 83.

In an alternative preferred embodiment, such analysis comprises the use of an oligonucleotide or oligomer for detecting the cytosine methylation state within genomic or treated (chemically modified) DNA, according to SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83. Said oligonucleotide or oligomer comprising a nucleic acid sequence having a length of at least nine (9) nucleotides which hybridizes, under moderately stringent or stringent conditions (as defined herein above), to a treated nucleic acid sequence according to SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83 and/or sequences complementary thereto, or to a genomic sequence according to SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 and SEQ ID NO: 64 to SEQ ID NO: 67 and/or

sequences complementary thereto.

Thus, the present invention includes nucleic acid molecules (e.g., oligonucleotides and peptide nucleic acid (PNA) molecules (PNA-oligomers)) that hybridise under moderately stringent and/or stringent hybridisation conditions to all or a portion of the sequences SEQ ID NO: 1 to SEQ ID NO: 55 and SEQ ID NO: 64 to SEQ ID NO: 83, or to the complements thereof. The hybridising portion of the hybridizing nucleic acids is typically at least 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or 35 nucleotides in length. However, even longer molecules have inventive utility, and are thus within the scope of the present invention.

Preferably, the hybridising portion of the inventive hybridising nucleic acids is at least 95%, or at least 98%, or 100% identical to the sequence, or to a portion thereof of SEQ ID NO: 1 to SEQ ID NO: 55 and SEQ ID NO: 64 to SEQ ID NO: 83, or to the complements thereof.

Hybridising nucleic acids of the type described herein can be used, for example, as a primer (e.g., a PCR primer), or a diagnostic and/or prognostic probe or primer. Preferably, hybridisation of the oligonucleotide probe to a nucleic acid sample is performed under stringent conditions and the probe is 100% identical to the target sequence. Nucleic acid duplex or hybrid stability is expressed as the melting temperature or Tm, which is the temperature at which a probe dissociates from a target DNA. This melting temperature is used to define the required stringency conditions.

In a preferred embodiment the present invention includes nucleic acid molecules comprising in each case a contiguous sequence of at least 9 nucleotides that is complementary to, or hybridizes under moderately stringent or stringent conditions to a sequence selected from the group consisting of SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83 and complements thereof, wherein the nucleic acid molecule comprises at least one TpA or CpA dinucleotide at a position where the corresponding untreated nucleic acid molecule according to SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 comprises a CpG dinucleotide. Preferred lengths of such nucleic acid molecules are 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35 nucleotides.

For target sequences that are related and substantially identical to the corresponding sequence of SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 and SEQ ID

NO: 64 to SEQ ID NO: 67 (such as allelic variants and SNPs), rather than identical, it is useful to first establish the lowest temperature at which only homologous hybridisation occurs with a particular concentration of salt (e.g., SSC or SSPE).Then, assuming that 1% mismatching results in a 1°C decrease in the Tm, the temperature of the final wash in the hybridisation reaction is reduced accordingly (for example, if sequences having > 95% identity with the probe are sought, the final wash temperature is decreased by 5°C).In practice, the change in Tm can be between 0.5°C and 1.5°C per 1% mismatch.

Examples of inventive oligonucleotides of length X (in nucleotides), as indicated by polynucleotide positions with reference to, e.g., SEQ ID NO:1, include those corresponding to sets (sense and antisense sets) of consecutively overlapping oligonucleotides of length X, where the oligonucleotides within each consecutively overlapping set (corresponding to a given X value) are defined as the finite set of Z oligonucleotides from nucleotide positions:
n to (n + (X-1));
where n=1, 2, 3,...(Y-(X-1));
where Y equals the length (nucleotides or base pairs) of SEQ ID NO: 1 (3190);
where X equals the common length (in nucleotides) of each oligonucleotide in the set (e.g., X=20 for a set of consecutively overlapping 20-mers); and
where the number (Z) of consecutively overlapping oligomers of length X for a given SEQ ID NO of length Y is equal to Y-(X-1).For example Z= 3190-19= 3171 for either sense or antisense sets of SEQ ID NO: 1, where X=20.

Preferably, the set is limited to those oligomers that comprise at least one CpG, TpG or CpA dinucleotide.

Examples of inventive 20-mer oligonucleotides include the following set of oligomers (and the antisense set complementary thereto), indicated by polynucleotide positions with reference to SEQ ID NO: 1: 1-20, 2-21, 3-22, 4-23, 5-24, ...3171-3190.

Preferably, the set is limited to those oligomers that comprise at least one CpG, TpG or CpA dinucleotide.

Likewise, examples of inventive 25-mer oligonucleotides include the following set of oligomers (and the antisense set complementary thereto), indicated by polynucleotide positions with reference to SEQ ID NO: 1: 1-25, 2-26, 3-27, 4-28, 5-29, ...3164-3190.

Preferably, the set is limited to those oligomers that comprise at least one CpG, TpG or CpA dinucleotide.

The present invention encompasses, for each of SEQ ID NO: 1 to SEQ ID NO: 55 and SEQ ID NO: 64 to SEQ ID NO: 83 (sense and antisense), multiple consecutively overlapping sets of oligonucleotides or modified oligonucleotides of length X, where, e.g., X= 9, 10, 17, 20, 22, 23, 25, 27, 30 or 35 nucleotides.

The oligonucleotides or oligomers according to the present invention constitute effective tools useful to ascertain genetic and epigenetic parameters of the genomic sequence corresponding to SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 and SEQ ID NO: 64 to SEQ ID NO: 67.

Particularly preferred oligonucleotides or oligomers according to the present invention are those in which the cytosine of the CpG dinucleotide (or of the corresponding converted TpG or CpA dinculeotide) sequences is within the middle third of the oligonucleotide; that is, where the oligonucleotide is, for example, 13 bases in length, the CpG, TpG or CpA dinucleotide is positioned within the fifth to ninth nucleotide from the 5'-end.

The oligonucleotides of the invention can also be modified by chemically linking the oligonucleotide to one or more moieties or conjugates to enhance the activity, stability or detection of the oligonucleotide. Such moieties or conjugates include chromophores, fluorophores, lipids such as cholesterol, cholic acid, thioether, aliphatic chains, phospholipids, polyamines, polyethylene glycol (PEG), palmityl moieties, and others. The probes may also exist in the form of a PNA (peptide nucleic acid) which has particularly preferred pairing properties. Thus, the oligonucleotide may include other appended groups such as peptides, and may include hybridisation-triggered cleavage agents (Krol et al., BioTechniques 6:958-976, 1988) or intercalating agents (Zon, Pharm. Res. 5:539-549, 1988). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a chromophore, fluorophor, peptide, hybridisation-triggered cross-linking agent, transport agent, hybridisation-triggered cleavage agent, etc.

The oligonucleotide may also comprise at least one art-recognised modified sugar and/or base moiety, or may comprise a modified backbone or non-natural internucleoside linkage.

The oligonucleotides or oligomers according to particular embodiments of the present invention are typically used in 'sets,' which contain at least one oligomer for analysis of each of the CpG dinucleotides of genomic sequences SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 and sequences complementary thereto, or to the corresponding CpG, TpG or CpA dinucleotide within a sequence of the treated nucleic acids according to SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83 and sequences complementary thereto. However, it is anticipated that for economic or other factors it may be preferable to analyse a limited selection of the CpG dinucleotides within said sequences, and the content of the set of oligonucleotides is altered accordingly.

Therefore, in particular embodiments, the present invention provides a set of at least two (2) (oligonucleotides and/or PNA-oligomers) useful for detecting the cytosine methylation state in treated genomic DNA (SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83), or in genomic DNA (SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 and sequences complementary thereto). These probes enable diagnosis, and/or classification of genetic and epigenetic parameters of lung cell proliferative disorders. The set of oligomers may also be used for detecting single nucleotide polymorphisms (SNPs) in treated genomic DNA (SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83), or in genomic DNA (SEQ ID NO: 1 to SEQ ID NO: 11 and SEQ ID NO: 64 to SEQ ID NO: 67 and sequences complementary thereto).

In preferred embodiments, at least one, and more preferably all members of a set of oligonucleotides is bound to a solid phase.

In further embodiments, the present invention provides a set of at least two (2) oligonucleotides that are used as 'primer' oligonucleotides for amplifying DNA sequences of one of SEQ ID NO: 1 to SEQ ID NO: 55 and SEQ ID NO: 64 to SEQ ID NO: 83 and sequences complementary thereto, or segments thereof.

It is anticipated that the oligonucleotides may constitute all or part of an "array" or "DNA

chip" (i.e., an arrangement of different oligonucleotides and/or PNA-oligomers bound to a solid phase). Such an array of different oligonucleotide- and/or PNA-oligomer sequences can be characterised, for example, in that it is arranged on the solid phase in the form of a rectangular or hexagonal lattice. The solid-phase surface may be composed of silicon, glass, polystyrene, aluminium, steel, iron, copper, nickel, silver, or gold. Nitrocellulose as well as plastics such as nylon, which can exist in the form of pellets or also as resin matrices, may also be used. An overview of the state of the art in oligomer array manufacturing can be gathered from a special edition of Nature Genetics (Nature Genetics Supplement, Volume 21, January 1999, and from the literature cited therein). Fluorescently labelled probes are often used for the scanning of immobilised DNA arrays. The simple attachment of Cy3 and Cy5 dyes to the 5'-OH of the specific probe are particularly suitable for fluorescence labels. The detection of the fluorescence of the hybridised probes may be carried out, for example, via a confocal microscope. Cy3 and Cy5 dyes, besides many others, are commercially available.

It is also anticipated that the oligonucleotides, or particular sequences thereof, may constitute all or part of an "virtual array" wherein the oligonucleotides, or particular sequences thereof, are used, for example, as 'specifiers' as part of, or in combination with a diverse population of unique labelled probes to analyse a complex mixture of analytes. Such a method, for example is described in US 2003/0013091 (United States serial number 09/898,743, published 16 January 2003). In such methods, enough labels are generated so that each nucleic acid in the complex mixture (i.e., each analyte) can be uniquely bound by a unique label and thus detected (each label is directly counted, resulting in a digital read-out of each molecular species in the mixture). In the final step of the method the determination of the presence, absence, classification or prognosis is determined according to the methylation status of the analyzed CpG positions. This is carried out by reference to a pre-existing data set that defines the methylation patterns characteristic to each disease phenotype. Preferably, the correlation of the methylation status of the marker CpG positions with the phenotypic parameters is done substantially without human intervention. Machine learning algorithms automatically analyse experimental data, discover systematic structure in it, and distinguish relevant parameters from uninformative ones.

Machine learning predictors are trained on the methylation patterns at the investigated CpG sites of the samples with known phenotypical classification. The CpG positions which prove to be discriminative are used to define the reference data set. This method is successful in

cancer classification (Model, F., Adorjan, P., Olek, A., and Piepenbrock, C., *Bioinformatics*. 17 Suppl 1:157-164, 2001).

A further aspect of the present invention is a kit useful for detecting distinguishing between or among gynaecological cell proliferative disorders of a subject comprising:
At least one of a bisulfite reagent, or a methylation-sensitive restriction enzyme; and at least one nucleic acid molecule or peptide nucleic acid molecule comprising, in each case a contiguous sequence of at least 9 nucleotides that is complementary to, or hybridizes under moderately stringent or stringent conditions to a sequence selected from the group consisting of SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ Id NO: 68 to SEQ ID NO: 83, and complements thereof. The kit can further comprise standard reagents for performing a methylation assay selected from the group consisting of MSP, MethyLight®, HeavyMethyl and combinations thereof. Additional components can include at least two oligonucleotides, whose sequences in each case correspond, are complementary or hybridize under stringent or highly stringent conditions to a 16 bp-long segment of the sequences SEQ ID NO: 1to SEQ ID NO: 55 and SEQ ID NO: 64 to SEQ ID NO: 83, preferably SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83.

Moreover, an additional aspect of the present invention is a kit comprising, for example: a bisulfite-containing reagent; a set of primer oligonucleotides containing at least two oligonucleotides whose sequences in each case correspond, are complementary, or hybridize under stringent or highly stringent conditions to a 16-base long segment of the sequences SEQ ID NO: 1 to SEQ ID NO: 55 and SEQ ID NO: 64 to SEQ ID NO: 83 (most preferably SEQ ID NO: 12 to SEQ ID NO: 55 and SEQ ID NO: 68 to SEQ ID NO: 83); oligonucleotides and/or PNA-oligomers; as well as instructions for carrying out and evaluating the described method. In a further preferred embodiment, said kit may further comprise standard reagents for performing a CpG position-specific methylation analysis, wherein said analysis comprises one or more of the following techniques: MSP, MethyLight ™ and HeavyMethyl™. However, a kit along the lines of the present invention can also contain only part of the aforementioned components.

Typical reagents (*e.g.*, as might be found in a typical MethyLight®-based kit) for MethyLight® analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); TaqMan® probes; optimized PCR buffers

and deoxynucleotides; and Taq polymerase.

Typical reagents (*e.g.*, as might be found in a typical MSP-based kit) for MSP analysis may include, but are not limited to: methylated and unmethylated PCR primers for specific gene (or methylation-altered DNA sequence or CpG island), optimized PCR buffers and deoxynucleotides, and specific probes.

Typical reagents (*e.g.*, as might be found in a typical HeavyMethyl®-based kit) for HeavyMethyl® analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); HeavyMethy blocking oligonucleotides; optimized PCR buffers and deoxynucleotides; and polymerase.

Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kits (*e.g.*, precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

The present invention shall now be further described without limitation thereof in the following examples with respect to the accompanying Figures and the attached sequence protocol. All publications cited herein are herewith incorporated by reference.

In the Figures,

Fig. 1 shows a flowchart of procedures for tampon insertion and sample preparation.
A - Tampon insertion.
B - Tampon transfer into a 50-ml tube.
C - Addition of 1.2 ml PBS buffer onto the tampon.
D - Centrifugation at 1000g for 10 min.
E – (1) shows the supernatant and (2) shows the pellet.
F - 0.2-ml aliquots of the supernatant fraction were each mixed with 0.2 ml of the working solution of the HighPure Viral Nucleic Acid kit and stored at –30°C until DNA isolation.

Fig. 2 shows the PMR values of 38 genes in patients without endometrial cancer (N=4) and with endometrial cancer (N=5) the genes *RASSF1A*, *hMLH1*, *CDH13*, *HSPA2* and *SOCS2* are highlighted with arrows.

Fig. 3 shows the methylation status of the five investigated genes (full box = methylated, empty box = unmethylated), status (endometrial cancer or no endometrial cancer) and age. * indicates cervical intraepithelial neoplasia grade III (CIN III); †indicates invasive cervical cancer.

Fig. 4 shows the ROC for detection of endometrial cancer by methylation analysis of DNA obtained from vaginal secretion for patients between 50 and 75 years of age (excluding patients with CIN III or cervical cancer, area under the curve is 0.988.

Fig. 5 shows the trend to increasing methylation from low-grade SIL to invasive cervical cancer. A statistically significant correlation between type of histology and number of methylated genes (ρ = 0. 47, P = 0.001) was found. No case had 10 or 11 methylated genes. Figure 5B shows High grade SILs (n=31). Figure 5A shows low grade/no dysplasia SILs (n=13) and fgure 5C shows invasive cervical cancer (n=5). The X axis scale shows the percentage of methylation positive samples, the Y axis shows the number of methylated genes.

Fig. 6 shows DNA methylation status of 11 genes in 49 cervicovaginal specimens was analysed using the MethyLight technique. A gene was deemed methylated if the PMR value was > 0 (Material and Methods; white and pink indicate unmethylated and methylated, respectively). Groups of patients were determined using unsupervised agglomerative hierarchical cluster analysis (average linkage, Manhattan distance) to group specimens and CpG regions. Gene names are given at the top of the figure. The inventors observed two clusters: 1 (red), 2 (blue). All patients with cervical cancer are grouped together in one cluster (blue).

Fig. 7 shows disease-free (A) and overall survival (B) according to *CDH1/CDH13* methylation status in serum samples. The lower dotted lines represent the methylated samples, the unbroken upper lines represent the unmethylated samples.

Figure 8 shows the polymerase mediated amplification of a CpG-rich sequence using methylation specific primers on four representative bisulfite-treated DNA strands (example cases "A"-"D") ("MSP Amplification"). The methylation specific forward and reverse primers ("1"), in each case, can anneal to the bisulfite-treated DNA strand ("3") if the corresponding subject genomic CpG sequences were methylated. The bisulfite-treated DNA strand ("3") can

be amplified if both forward and reverse primers ("1") anneal, as shown in representative case "A" at the top of the figure.

Figure 9 shows polymerase-mediated amplification analysis of bisulfite-treated DNA ("3") corresponding to a CpG-rich genomic sequence by means of the MethylHeavy® technique. Amplification of the treated DNA ("3") is precluded if the blocking oligonucleotide ("5") anneals to the treated DNA as shown for the example case "B."

Examples

Example 1

The following study was performed to determine whether it is possible to detect endometrial cancer by analyzing methylated DNA in cervicovaginal secretion.

Patients and Samples. A total of 124 patients were recruited for this study: 15 patients had endometrial cancer, while the no endometrial cancer group contained five patients with invasive cervical cancer, 35 with cervical intraepithelial neoplasia (CIN I, three cases; CIN II, 19 cases; CIN III, 13 cases), and 69 patients with benign disease of the uterus. Sample collection was done between 01.01.2003 and 31.05.2003 at the Department of Obstetrics and Gynecology, Innsbruck University Hospital, Austria. All patients who were scheduled to undergo surgery of the uterus on the next day including a histological diagnosis were invited to attend the study. Samples and clinical data were collected after informed consent was obtained. To ensure standardized sample collection a tampon was inserted in the patient by a physician after speculum examination and retained intravaginal for 30 minutes. Preparation of the samples is shown in Figure 1.

DNA Isolation and Methylation Analysis. Genomic DNA from samples was isolated using the *High Pure Viral Nucleic Acid Kit* (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's protocol with some modifications for multiple loading of the DNA extraction columns to gain a sufficient amount of DNA. Sodium bisulfite-treated genomic DNA was analyzed by means of MethyLight, a fluorescence-based, real-time PCR assay, as described previously. Briefly, two sets of primers and probes, designed specifically for bisulfite-converted DNA, were used: a methylated set for the gene of interest and a reference set, ß-actin (*ACTB*), to normalize for input DNA. Specificity of the reactions for methylated DNA

was confirmed separately using *Sss*I (New England Biolabs)-treated human white blood cell DNA (heavily methylated). The percentage of fully methylated molecules at a specific locus was calculated by dividing the *GENE:ACTB* ratio of a sample by the *GENE:ACTB* ratio of *Sss*I-treated white blood cell DNA and multiplying by 100. The abbreviation PMR (percentage of fully methylated reference) indicates this measurement. A gene was deemed methylated if the PMR value was > 0. Primer and probes specific for methylated DNA and used for MethyLight reactions are listed in Table 1.

Reaction Conditions. The reactions were run with the following assay conditions: *Reaction solution* comprised: 300 nM forward primer; 300 nM reverse primer and 100 nM probe; plus Magnesium Chloride; taq polymerase, dNTPs andDNA, in a final reaction volume of 30 μl); *Cycling conditions*: 95°C for 10 minutes; then 50 cycles of: 95°C for 15 seconds; 60°C for 1 minute.

Statistical Analysis. Associations between categorical variables where tested using the Chi-Square test. Differences in median of age were examined with the Mann-Whitney U test or – between more than two groups – with the Kruskal Wallis test. Due to a significant age difference between endometrial cancer patients and the no endometrial cancer group, an age-matched group of the no endometrial cancer group with a matching ratio 1:2 was randomly selected. The computation of the Matching-Group was done using MATLAB R12 (www.mathworks.com). For determination of diagnostic accuracy a non-parametric Receiver Operating Curve (ROC) Analysis with Linear Interpolation was performed. A P-value of less than 0.05 was considered statistically significant. All statistical calculations were performed using SPSS, version 11.0, for Windows.

Results

Aberrant methylation of 38 genes in DNA obtained from vaginal secretion from the first five patients with endometrial cancer and the first four patients with benign disease was analyzed to determine appropriate genes for further study. The most appropriate genes for our further analyses were determined to be those that revealed the greatest difference in PMR values between patients with benign disease of the uterus and endometrial cancer patients (Fig. 2). Five genes, namely *RASSF1A*, *hMLH1*, *CDH13*, *HSPA2* and *SOCS2*, were selected for further analysis. DNA methylation in three or more of these five genes was observed in cervicovaginal secretion of all five patients with endometrial cancer, whereas all four patients without

endometrial cancer showed no or fewer than three genes to be methylated. The inventors therefore determined the cut-off value between no endometrial cancer and endometrial cancer as methylation positive in three or more of the five investigated genes.

The overwhelming majority of the patients without endometrial cancer (99 of 109) revealed no or fewer than three genes methylated, whereas all of the 15 endometrial cancer patients had three or more genes methylated in their vaginal secretion (Fig. 3, $P<0.001$, chi^2 test). Histological examination of the ten patients in the no endometrial cancer group with three or more genes methylated revealed invasive cervical cancer (four cases), CIN III (one case), endometrium polyp (four cases) and fibroids (one case). Samples were collected after primary surgery (curettage, punch biopsy of the cervix or hysteroscopic operation) and before secondary surgery (hysterectomy) in 16 out of 124 patients; 9/16 patients had endometrial cancer, 3/16 CIN III and 4/16 benign disease of the endometrium. All nine endometrial cancer patients had three or more genes methylated, the three CIN III patients revealed no methylated genes and one of the patients with benign disease showed one gene to be methylated. Within the group of patients from whom the vaginal secretion was collected prior to any surgery, one patient presented due to sonographically detected serometra with complete stenosis of the cervicouterine canal. Even this patient showed methylation of three of the five tested genes. DNA methylation of the five genes identified, seems to increase with age although statistically not significant (data not shown). Using all 15 endometrial cancer cases and 109 controls, the area under the ROC curve was 0.973 (data not shown). To rule out the possibility that abnormal DNA methylation is merely a surrogate for age rather than a cancer-specific marker, the inventors randomly age-matched two non-endometrial cancer controls for each endometrial cancer case. Investigation of DNA methylation in the cervicovaginal secretion of these 45 patients was still able to discriminate between endometrial cancer and patients without endometrial cancer ($P<0.001$, chi^2 test) with a sensitivity of 100% and a specificity of 80%.

When analyzing all patients between 50 years and 75 years of age and excluding patients with CIN III or cervical cancer sensitivity was 100% and the specificity rose to 97.2% (Fig. 4). In this group only one out of 35 samples was false positive.

n our study all endometrial cancer patients revealed three or more of the five investigated genes methylated, whereas 99 out of 109 patients without endometrial cancer had no or fewer than three genes methylated. Four out of ten patients in the no endometrial cancer group with

three or more genes methylated had invasive cervical cancer. These cases indicate that some cervical cancer patients can also be identified with this assay.

In some cases (16 out of 124) samples were collected after primary surgery and before secondary surgery. All endometrial cancer patients in this group had three or more genes methylated. These results demonstrate that aberrant methylation analysis can detect endometrial cancer even after primary surgery.

As endometrial cancer is more prevalent in older women and abnormal DNA methylation in non-malignant tissues seems to increase with age (18), the inventors especially addressed this problem within this project. Comparison of DNA methylation in the cervicovaginal secretion of endometrial cancer patients and age-matched non-endometrial cancer controls revealed still highly significant differences between these two groups.

Endometrial cancer occurs in almost all cases after menopause. Therefore the inventors analyzed all patients between 50 years and 75 years of age and excluded patients with CIN III or invasive cervical cancer. These patients represent the group that will benefit from an endometrial cancer screening assay. In these group the sensitivity and specificity to detect patients with endometrial cancer was 100% and 97.2%, respectively.

Long-term tamoxifen users who are at increased risk for endometrial cancer have a worse prognosis for such cancers, which seems to be due to less favorable histology and higher stage. This indicates an urgent need for a simple, non-invasive means of early detection of endometrial cancer, especially in this subgroup of women.

Example 2

The aim of this study was to examine whether HPV DNA testing and methylation analysis of DNA obtained from cervicovaginal specimens, collected on a tampon, are able to detect invasive cervical cancer and whether these changes are already present in the precursor lesions.

Patients and Samples. A total of 34 patients with cervical intraepithelial neoplasia and five patients with invasive cervical cancer were included in this study. Patients were referred to our hospital for further treatment because of abnormal PAP smear or obvious cervical lesion. Patients underwent cervical conisation or in the case of obvious carcinomatous lesion cervical

biopsy to obtain a histological diagnosis. Additionally, ten patients without cervical dysplasia but who underwent hysterectomy because of fibroids were investigated. All patients were taken from a prior study performed to determine whether it is possible to detect endometrial cancer by analyzing methylated DNA in cervicovaginal specimens collected by tampon (Example 1).

Sample collection Samples and clinical data were collected after informed consent was obtained. To ensure standardized sample collection a tampon was inserted in the vagina by a physician after speculum examination prior to surgery (the day before) and retained intravaginal for 30 minutes. Preparation of the samples was as described in Example 1 and Figure 1. Briefly, the tampon was transferred to a 50-ml tube after removal and 1.2 ml PBS buffer was added to the tampon. Centrifugation at 1000g for 10min produced supernatant and a pellet. Aliquots (0.2ml) of the supernatant were each mixed with 0.2ml from the working solution of the High Pure Viral Nucleic Acid Kit (Roche Diagnostics, Mannheim, Germany) and stored at -30 C until DNA isolation.

DNA isolation and methylation analysis. Genomic DNA from samples was isolated using the *High Pure Viral Nucleic Acid Kit* (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's protocol with some modifications for multiple loading of the DNA extraction columns to gain a sufficient amount of DNA. After sodium bisulfite conversion, 11 genes (*SOCS1*, *CDH1*, *TIMP3*, *GSTP1*, *DAPK*, *hTERT*, *CDH13*, *HSPA2*, *MLH1*, *RASSF1A* and *SOCS2*) underwent methylation analysis by means of the fluorescence-based, real-time PCR MethyLight assay as described previously. Briefly, two sets of primers and probes, designed specifically for bisulfite-converted DNA, were used: a methylated set for the gene of interest and a reference set, ß-actin (*ACTB*), to normalize for input DNA. Specificity of the reactions for methylated DNA was confirmed separately using *Sss*I (New England Biolabs)-treated human white blood cell DNA (heavily methylated). The percentage of fully methylated molecules at a specific locus was calculated by dividing the *GENE:ACTB* ratio of a sample by the *GENE:ACTB* ratio of *Sss*I-treated DNA and multiplying by 100. The result is given as PMR (percentage of fully methylated reference). A gene was deemed methylated if the PMR value was > 0. Primers and probes used for MethyLight reactions are listed in Table 2.

MethyLight Reaction Conditions. The reactions were run with the following assay conditions: *Reaction solution* comprised: 300 nM forward primer; 300 nM reverse primer and 100

nM probe; plus magnesium chloride; taq polymerase, dNTPs andDNA, in a final reaction volume of 30 µl); *Cycling conditions*: 95°C for 10 minutes; then 50 cycles of: 95°C for 15 seconds; 60°C for 1 minute .

HPV DNA Analysis. HPV PCR enzyme immunoassay (PCR-EIA) was done as described by Jacobs, M. V., Snijders, P. J., van den Brule, A. J., Helmerhorst, T. J., Meijer, C. J., and Walboomers, J. M. A general primer GP5+/GP6(+)-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings. J. Clin. Microbiol., *35*: 791-795, 1997. Briefly, 10 µl of purified total cellular DNA was employed for PCR using consensus primers GP5+/bioGP6+. The reaction mixture contained 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 3.5 mM $MgCl_2$, 200 µM of each dNTP, 1 U of thermo-stable DNA polymerase (Taq DNA Polymerase, Roche, Vienna, Austria), and 50 pmol each of the GP5+ (5-TTTGTTACTGTGGTAGATATACTAC-3 SEQ ID NO:56) and bioGP6+ (5-GAAAAATAAACTGTAAATCATATT-3 SEQ ID NO:56) primers (MWG-Biotech, Ebersberg, Germany). A 4-min denaturation step at 94°C was followed by 40 cycles of amplification in a PCR cycler (Gene Amp PCR System 9600, Perkin Elmer, Norwalk, USA). To determine the HPV type, the PCR product was subjected to an enzyme immunoassay (EIA) recognizing 14 different high-risk HPV types as described by Jacobs et al.(as above).

Statistical Analysis. Associations between categorical variables were tested with Pearson's chi square test (or Fisher's exact test) and the Mantel-Haenzel Test. Correlations between ordinal variables were evaluated using the Spearman Rank Correlation Coefficient. For unsupervised hierarchical clustering of clinical cases and genes the inventors used the complete linkage aggregation method and the Manhattan distance function.

All statistical calculations were performed using SPSS, version 11.0, for Windows and Gene Expression Similarity Suite (http://genome.tugraz.at/Software/Genesis/Genesis.html).

RESULTS

This study investigated cervicovaginal specimens sampled by intravaginal tampon application in patients without cervical dysplasia, patients with low-grade and high-grade SIL and patients with invasive cervical cancer for HPV DNA and aberrant methylation of 11 genes. High-risk HPV DNA was detected in 2/3, 21/31, and 3/5 in low-grade SIL, high- grade SIL

and invasive cervical cancer, respectively (Table 3). Patients without cervical dysplasia showed no HPV infection. The three HPV-positive cervical cancers were large cell squamous cancers, whereas both HPV-negative cases were small cell cervical cancers. Differences in methylation of the investigated genes between the non-dysplasia group and the low-grade SIL group were not statistically significant. Therefore, the results from these two groups were combined and compared with the high-grade SIL and invasive cancer groups. An overview of the frequency of methylated genes is given in Table 3. All investigated genes except *GSTP1* and *SOCS2* were significantly more frequently methylated in high-grade SIL and/or invasive cancer in comparison to the non-dysplasia/low-grade SIL group. *CDH1* and *SOCS2* were found to be methylated in nearly a quarter of the non-dysplasia/low-grade SIL patients whereas *hTERT* was methylated exclusively in specimens obtained from cervical cancer patients. No methylated genes were found in 61% and 42% of the cervicovaginal specimens from patients with non-dysplasia/low-grade SIL and high-grade SIL, respectively, whereas cervical cancer patients revealed five or more methylated genes in each investigated sample. The percentage of methylation-positive samples showed a significantly increasing trend from non-dysplasia/low-grade SIL to invasive cervical cancer, as shown in Fig. 5. No significant correlation between HPV positivity and aberrant hypermethylation was observed ($p = 0.295$, Fisher's exact test). HPV DNA-positive samples and/or at least one methylated gene were found in 46% (7/13; non-dysplasia/low-grade SIL), 94% (29/31; high-grade SIL), and 100% (5/5, invasive cancer) of samples (Table 3).

Two clusters were formed by unsupervised hierarchical cluster analysis using solely information on the DNA methylation of the 11 genes tested. One of the two clusters contained all five invasive cancers as well as two high-grade SILs (Fig. 3).

hTERT was never found to be methylated in cervical intraepithelial neoplasias, whereas 80% of the specimens obtained from cervical cancer patients were methylated. This finding suggests that methylation of *hTERT* is a late event in cervical carcinogenesis. *SOCS2* and *CDH1* were methylated in nearly a quarter of patients from the non-dysplasia/low-grade SIL groups, indicating that methylation of these genes is an early event in cervical carcinogenesis. Our results clearly show that an increasing percentage of methylation-positive samples is associated with increasing pathological changes of the cervix uteri (Fig. 5), suggesting that methylation, in addition to HPV infection, is an important factor in cervical carcinogenesis. HPV DNA-positive samples and/or at least one methylated gene were found in more than 90% of

samples from patients with high-grade SIL and in 100% of patients with invasive cervical cancer. No significant correlation between HPV positivity and aberrant hypermethylation was observed. It can be speculated that aberrant methylation in women with or without HPV infection may help identify subgroups at increased risk for histological progression or cancer development. Our study shows that it is possible to detect HPV and aberrant hypermethylation of various genes in DNA from tampon-collected samples. The inventors were able to identify all invasive cervical cancers. Both methods combined detect a high percentage of high-grade cervical lesions.

Example 3

In the following study the inventors investigated the methylation status of *CDH1* and *CDH13* in serum samples of cervical cancer patients for their utility as prognostic markers.

Patients and Samples

A total of 93 patients with invasive cervical cancer (age 26-96 years, median 52 years), all treated at the Department of Obstetrics and Gynecology, Innsbruck University Hospital, between 1990 and 1998 were included in this study. Serum samples were taken on the date of diagnosis and before initial treatment. These serum samples were taken from a prior study investigating the presence of serum human papillomavirus DNA in cervical cancer patients.

Major clinical and histopathological characteristics of patients are given in Table 4. Treatment was according to international standards. None of the patients received concurrent chemotherapy and radiotherapy. All patients were followed up after primary treatment at our department, namely at intervals increasing from three months to one year until death or end of the study. The follow-up period ranged from one month to 12.4 years (median 3.5 years).

DNA isolation and methylation analysis

Serum samples (300µl) were treated with SDS and proteinase K (300µl of 1% SDS, 500µg/ml proteinase K) at 55°C overnight, followed by phenol/chloroform extraction and ethanol precipitation of DNA. The DNA was re-suspended in 80µl LoTE buffer (30mM Tris and 0.3mM EDTA). Sodium bisulfite conversion of genomic DNA was performed as described previously.[12] Sodium bisulfit-treated genomic DNA was analysed by means of the MethyLight, a fluorescence-based, real-time PCR assay, as described previously. Briefly, two sets of primers and probes, designed specifically for bisulfite-converted DNA, were used: a methylated set

for the gene of interest and a reference set, ß-actin (*ACTB*), to normalize for input DNA. Specificity of the reactions for methylated DNA was confirmed separately using *Sss*I (New England Biolabs)-treated human white blood cell DNA (heavily methylated). The percentage of fully methylated molecules at a specific locus was calculated by dividing the *GENE:ACTB* ratio of a sample by the *GENE:ACTB* ratio of *Sss*I-treated white blood cell DNA and multiplying by 100. The abbreviation PMR (percentage of fully methylated reference) indicates this measurement. For each MethyLight reaction 10µl of bisulfite-treated genomic DNA were used. A gene was deemed methylated if the PMR value was > 0. To verify the reproducibility of each assay the normalized value (*Gene*: *ACTB*) of the standard sample was compared between the different PCR runs. The following primers and probes were used for MethyLight reactions: *CDH1*: 5'-AATTTTAGGTTAGAGGGTTATCGCGT-3' SEQ ID NO: 58 (forward primer), 5'-TCCCCAAAACGAAACTAACGAC-3' SEQ ID NO: 59 (reverse primer), 5'-FAM-CGCCCACCCGACCTCGCAT-BHQ-1-3' SEQ ID NO: 60 (probe); *CDH13*: 5'-AATTTCGTTCGTTTTGTGCGT-3' SEQ ID NO: 61 (forward primer), 5'-CTACCCGTACCGAACGATCC-3' SEQ ID NO: 62 (reverse primer), 5'-FAM-AACGCAAAACGCGCCCGACA-BHQ-1-3' SEQ ID NO: 63 (probe).

MethyLight Reaction Conditions. The reactions were run with the following assay conditions: *Reaction solution* comprised: 300 nM forward primer; 300 nM reverse primer and 100 nM probe; plus magnesium chloride; taq polymerase, dNTPs and DNA, in a final reaction volume of 30 µl); *Cycling conditions*: 95°C for 10 minutes; then 50 cycles of: 95°C for 15 seconds; 60°C for 1 minute .

Statistical analysis

Associations between categorical variables were tested with Pearson's chi square test. The Kaplan-Meier method was used for univariate survival analysis, and the log rank test was used to assess the difference between survival curves. Cox's proportional hazards analysis was used to estimate the prognostic effects of various variables. A P value of less then 0.05 was considered statistically significant. These statistical calculations were performed using SPSS, version 11.0, for Windows.

RESULTS

Aberrant promoter hypermethylation of *CDH1* and *CDH13* was observed in 42% (39 out of 93) and 4% (4 out of 93), respectively (Table 4). Three of the *CDH13* methylation positive

serum samples also revealed *CDH1* methylation. The inventors therefore collapsed *CDH1* and/or *CDH13* methylation for further analysis. *CDH1* methylation was predominantly observed in FIGO stage III, whereas distribution of *CDH1* and *CDH13* methylation within the other clinical and histopathological parameters showed no significant differences (Table 4).

To determine whether any prognostic significance was attached to differences in *CDH1* and/or *CDH13* methylation, the inventors compared the clinical outcome of cervical cancer patients with *CDH1/CDH13* methylation status. A trend to poorer overall survival for patients with methylated *CDH1/CDH13* was observed (P = 0.09) (Figure 7B). Cervical cancer patients with unmethylated *CDH1/CDH13* in serum samples taken before treatment revealed a statistically significant better disease-free survival in comparison to patients with methylated *CDH1/CDH13* (P = 0.03) (Figure 7A). Median disease-free survival for *CDH1/CDH13* methylation negative and positive patients was 4.3 years and 1.2 years, respectively.

To assess independent prognostic significance a Cox proportional hazard model analysis was performed including tumor stage, histology, grade of differentiation, age and *CDH1/CDH13* methylation status. In addition to tumor stage and age only *CDH1/CDH13* methylation status (P = 0.005) turned out to be of independent prognostic significance for disease-free and overall survival in cervical cancer patients (Tables 5 and 6). Serum *CDH1/CDH13* methylation positive patients had a more than twofold risk for relapse and death than did *CDH1/CDH13* methylation negative patients.

Distribution of *CDH1* and *CDH13* methylation within the clinical and histopathological parameters showed no significant differences except *CDH1* methylation was predominantly observed in advanced FIGO stage. Therefore, no difference in *CDH1* methylation according to FIGO stage could be observed. The higher methylation frequency of CDH1 with increasing tumor stage and the association of *CDH1* and *CDH13* methylation in serum samples with enhanced relapse frequency in our study confirms these results. Up to now it has not been investigated whether one of the above mentioned mechanisms alone or a combination of these mechanisms causes loss of cadherin expression. DNA recovered from plasma or serum of patients with various malignancies reflects tumor-specific genetic and epigenetic alterations like methylation of the primary tumor (Ziegler A, Zangemeister-Wittke U, Stahel RA. Circulating DNA: a new diagnostic gold mine? Cancer Treat Rev 2002;28:255-71.). The simple procedure of blood drawing in combination with a high-throughput analysis like MethyLight

opens a feasible approach for a possible routine use of these markers. Inactivation of the cadherin-mediated cell adhesion system, caused by aberrant methylation, is a common finding in human cancers. Therefore, investigation of *CDH1* and *CDH13* methylation as a prognostic parameter in serum samples from patients with various malignancies could be of interest. Our study revealed that *CDH1/CDH13* methylation is an independent prognostic parameter for both disease free and overall survival in cervical cancer patients.

Example 4

The following study was performed to determine whether it is possible to detect endometrial cancer, ovarian cancer and invasive cervical cancer by analyzing methylated DNA in cervicovaginal secretion of the genes SFRP2, SFRP4, SFRP5 and CCND2.

Patients and Samples. A total of 24 patients were recruited for this study: 4 patients had endometrial cancer, 2 patients had ovarian cancer and 6 patients had cervical cancer. To ensure standardized sample collection a tampon was inserted in the patient by a physician after speculum examination and retained intravaginal for 30 minutes. Preparation of the samples is shown in Figure 1.

DNA Isolation and Methylation Analysis. Genomic DNA from samples was isolated using the *High Pure Viral Nucleic Acid Kit* (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's protocol with some modifications for multiple loading of the DNA extraction columns to gain a sufficient amount of DNA. Sodium bisulfite-treated genomic DNA was analyzed by means of MethyLight, a fluorescence-based, real-time PCR assay, as described previously. Briefly, two sets of primers and probes, designed specifically for bisulfite-converted DNA, were used: a methylated set for the gene of interest and a reference set, ß-actin (*ACTB*), to normalize for input DNA. Specificity of the reactions for methylated DNA was confirmed separately using *Sss*I (New England Biolabs)-treated human white blood cell DNA (heavily methylated). The percentage of fully methylated molecules at a specific locus was calculated by dividing the *GENE:ACTB* ratio of a sample by the *GENE:ACTB* ratio of *Sss*I-treated white blood cell DNA and multiplying by 100. The abbreviation PMR (percentage of fully methylated reference) indicates this measurement. A gene was deemed methylated if the PMR value was > 0. Primer and probes specific for methylated DNA and used for MethyLight reactions are listed in Table 8.

MethyLight Reaction Conditions. The reactions were run with the following assay conditions: *Reaction solution* comprised: 300 nM forward primer; 300 nM reverse primer and 100 nM probe; plus magnesium chloride; taq polymerase, dNTPs and DNA, in a final reaction volume of 30 µl); *Cycling conditions*: 95°C for 10 minutes; then 50 cycles of: 95°C for 15 seconds; 60°C for 1 minute .

Statistical Analysis. Associations between categorical variables where tested using the Chi-Square test. Differences in median of age were examined with the Mann-Whitney U test or – between more than two groups – with the Kruskal Wallis test. Due to a significant age difference between endometrial cancer patients and the no endometrial cancer group, an age-matched group of the no endometrial cancer group with a matching ratio 1:2 was randomly selected. The computation of the Matching-Group was done using MATLAB R12 (www.mathworks.com). For determination of diagnostic accuracy a non-parametric Receiver Operating Curve (ROC) Analysis with Linear Interpolation was performed. A P-value of less than 0.05 was considered statistically significant. All statistical calculations were performed using SPSS, version 11.0, for Windows.

Results

See Table 9 for results. All ovarian cancer patients showed methylation of at least one gene of the panel SFRP2, SFRP4, SFRP5 and CCND2. All cervical cancer patients showed methylation of at least two genes of the panel SFRP2, SFRP4, SFRP5 and CCND2, with two patients exhibiting methylation of all genes. All endometrial cancer patients showed methylation of at least two genes of the panel SFRP2, SFRP4, SFRP5 and CCND2.

Table 1: Primers and probes used according to Example 1

HUGO Gene Nomenclature	Forward Primer Sequence	Reverse Primer Sequence	Probe Oligo Sequence
ACTB	TGGTGATGGAGGAGGTTTAGTAAGT	AACCAATAAAACCTACTCCTCCCTTAA	6FAM-ACCACCACCCAACACACAATAACAAACACA-BHQ-1
APC	GAACCAAAACGCTCCCCAT	TTATATGTCGGTTACGTGCGTTTATAT	6FAM-CCCGTCGAAAACCCGCCGATTA-BHQ-1
ARHI	GCGTAAGCGGAATTTATGTTTGT	CCGCGATTTTATATTCCGACTT	6FAM-CGCACAAAAACGAAATACGAAAACGCAAA-BHQ-1
BLT1	GCGTTGGTTTTATCGGAAGG	AAACCGTAATTCCCGCTCG	6FAM-GACTCCGCCCAACTTCGCCAAAA-BHQ-1
BRCA1	GAGAGGTTGTTGTTTAGCGGTAGTT	CGCGCAATCGCAATTTTAAT	6FAM-CCGCGCTTTTCCGTTACCACGA-BHQ-1
CALCA	GTTTTGGAAGTATGAGGGTGACG	TTCCCGCCGCTATAAATCG	6FAM-ATTCCGCCAATACACAACAACCAATAAACG-BHQ-1
CDH1	AATTTTAGGTTAGAGGGTTATCGCGT	TCCCCAAAACGAAACTAACGAC	6FAM-CGCCCACCCGACCTCGCAT-BHQ-1
CDH13	AATTTCGTTCGTTTTGTGCGT	CTACCCGTACCGAACGATCC	6FAM-AACGCAAAACGCGCCCGACA-BHQ-1
CDKN2A	TGGAGTTTTCGGTTGATTGGTT	AACAACGCCCGCACCTCCT	6FAM-ACCCGACCCCGAACCGCG-BHQ-1
CYP1B1	GTGCGTTTGGACGGGAGTT	AACGCGACCTAACAAAACGAA	6FAM-CGCCGCACACCAAACCGCTT-BHQ-1
DAPK1	TCGTCGTCGTTTCGGTTAGTT	TCCCTCCGAAACGCTATCG	6FAM-CGACCATAAACGCCAACGCCG-BHQ-1
ESR1	GGCGTTCGTTTTGGGATTG	GCCGACACGCGAACTCTAA,	6FAM-CGATAAAACCGAACGACCCGACGA-BHQ-1
ESR2	TTTGAAATTTGTAGGGCGAAGAGTAG	ACCCGTCGCAACTCGAATAA	6FAM-CCGACCCAACGCTCGCCG-BHQ-1
FGF18	ATCTCCTCCTCCGCGTCTCT	TCGCGCGTAGAAAACGTTT	6FAM-CGACCGTACGCATCGCCGC-BHQ-1
GSTM3	GCG CGA ACG CCC TAA CT	AAC GTC GGT ATT AGT CGC GTT T	6FAM-CCC CGT TCT CCG TCC CTT ACC TCC-BHQ-1
GSTP1	GTCGGCGTCGTGATTTAGTATTG	AAACTACGACGACGAAACTCCAA	6FAM-AAACCTCGCGACCTCCGAACCTTATAAAA-BHQ-1
HIC1	GTTAGGCGGTTAGGGCGTC	CCGAACGCCTCCATCGTAT	6FAM-CAACATCGTCTACCCAACACACTCTCCTACG-BHQ-1
HLA-G	CAC CCC CAT ATA CGC GCT AA	GGT CGT TAC GTT TCG GGT AGT TTA	6FAM-CGC GCT CAC ACG CTC AAA AAC CT-BHQ1
HSD17B4	TATCGTTGAGGTTCGACGGG	TCCAACCTTCGCATACTCACC	6FAM-CCCGCGCCGATAACCAATACCA-BHQ-1
HSPA2	CAC GAA CAC TAC CAA CAA CTC AAC T	GGG AGC GGA TTG GGT TTG	6FAM-CCG CGC CCA ATT CCC GAT TCT-BHQ1
IGFBP2	CTC GCG CCG ACA AAT AAA TAC	CGG GAA GAG TAG GGA ATT TTT AGA GT	6FAM-ACG CCC GCT CGC CCA CCT-BHQ1
MGMT	GCGTTTCGACGTTCGTAGGT	CACTCTTCCGAAAACGAAACG	6FAM-CGCAAACGATACGCACCGCGA-BHQ-1
MLH1	AGGAAGAGCGGATAGCGATTT	TCTTCGTCCCTCCCTAAAACG	6FAM-CCCGCTACCTAAAAAAATATACGCTTACGCG-BHQ-1
MLLT7	CCT CAC GAT ACC TCC CCT CAA	TTA GGG ATT AGC GTT TTG GGA TT	6FAM-AAA CAC ATT CCT ACC AAT CTT CAA AAA ATC GCG-BHQ1
MT3	CGA TAA ACG AAC TTC TCC AAA CAA	GCG CGG TGC GTA GGG	6FAM-AAA CGC GCG ACT TAA CTA ATA ACA ACA AAT AAC GA-BHQ-1
MYOD1	GAGCGCGCGTAGTTAGCG	TCCGACACGCCCTTTCC	6FAM-CTCCAACACCCGACTACTATATCCGCGAAA-BHQ-1
PGR	TTATAATTCGAGGCGGTTAGTGTTT	TCGAACTTCTACTAACTCCGTACTACGA	6FAM-ATCATCTCCGAAAATCTCAAATCCCAATAATACG-BHQ-1
PPP1R13B	CCT CAC CCA CCG ACA TCA TC	TCG GAG CGG TGG GTA TAG TTC	6FAM-AAA AAT CCG CGA CGC CCT CGA-BHQ-1
PTGS2	CGGAAGCGTTCGGGTAAAG	AATTCCACCGCCCCAAAC	6FAM-TTTCCGCCAAATATCTTTTCTTCTTCGCA-BHQ-1
RASSF1A	ATTGAGTTGCGGGAGTTGGT	ACACGCTCCAACCGAATACG	6FAM-CCCTTCCCAACGCGCCCA-BHQ-1
SOCS1	GCGTCGAGTTCGTGGGTATTT	CCGAAACCATCTTCACGCTAA	6FAM-ACAATTCCGCTAACGACTATCGCGCA-BHQ-1
SOCS2	TCC CTT CCC CGC CAT T	TTG TTT TTG TCG CGG TGA TTT	6FAM-CCG AAA AAC TCA AAA CAC CGC AAA ATC AT-BHQ1
SYK	GGGCGCGATATTGGGAG	GCGACTCTTCCTCATTTTAAACAAC	6FAM-CCTTAACGCGCCCGAACAAACG-BHQ-1
TERT	GGATTCGCGGGTATAGACGTT	CGAAATCCGCGCGAAA	6FAM-CCCAATCCCTCCGCCACGTAAAA-BHQ-1
TFF1	TAAGGTTACGGTGGTTATTTCGTGA	ACCTTAATCCAAATCCTACTCATATCTAAAA	6FAM-CCCTCCCGCCAAAATAAATACTATACTCACTACAAAA-BHQ-1
TIMP3	GCGTCGGAGGTTAAGGTTGTT	CTCTCCAAAATTACCGTACGCG	6FAM-AACTCGCTCGCCCGCCGAA-BHQ-1
TITF1	CGA AAT AAA CCG AAT CCT CCT TAA	TGT TTT GTT GTT TTA GCG TTT ACG T	6FAM-CTC GCG TTT ATT TTA ACC CGA CGC CA-BHQ-1
TP53BP2	ACC CCC TAA CGC GAC TTT ATC	GTT CGA TTC GGG ATT AGT TGG T	6FAM-CGC TCG TAA CGA TCG AAA CTC CCT CCT-BHQ-1
TWIST	GTAGCGCGGCGAACGT	AAACGCAACGAATCATAACCAAC	6FAM-CCAACGCACCCAATCGCTAAACGA-BHQ-1

Table 2: Primers and probes used according to Example 2

Gene	Forward Primer	Reverse Primer	Probe
SOCS1	GCGTCGAGTTCGTGGGTATTT	CCGAAACCATCTTCACGCTAA	6FAM-ACAATTCCGCTAACGACTATCGCGCA-BHQ-1
CDH1	AATTTTAGGTTAGAGGGTTATCGCGT	TCCCCAAAACGAAACTAACGAC	6FAM-CGCCCACCCGACCTCGCAT-BHQ-1
TIMP3	GCGTCGGAGGTTAAGGTTGTT	CTCTCCAAAATTACCGTACGCG	6FAM-AACTCGCTCGCCCGCCGAA-BHQ-1
GSTP1	GTCGGCGTCGTGATTTAGTATTG	AAACTACGACGACGAAACTCCAA	6FAM-AAACCTCGCGACCTCCGAACCTTATAAAA-BHQ-1
DAPK1	TCGTCGTCGTTTCGGTTAGTT	TCCCTCCGAAACGCTATCG	6FAM-CGACCATAAACGCCAACGCCG-BHQ-1
hTERT	GGATTCGCGGGTATAGACGTT	CGAAATCCGCGCGAAA	6FAM-CCCAATCCCTCCGCCACGTAAAA-BHQ-1
CDH13	AATTTCGTTCGTTTTGTGCGT	CTACCCGTACCGAACGATCC	6FAM-AACGCAAAACGCGCCCGACA-BHQ-1
HSPA2	CACGAACACTACCAACAACTCAACT	GGGAGCGGATTGGGTTTG	6FAM-CCGCGCCCAATTCCCGATTCT-BHQ1
MLH1	AGGAAGAGCGGATAGCGATTT	TCTTCGTCCCTCCCTAAAACG	6FAM-CCCGCTACCTAAAAAAATATACGCTTACGCG-BHQ-1
RASSF1A	ATTGAGTTGCGGGAGTTGGT	ACACGCTCCAACCGAATACG	6FAM-CCCTTCCCAACGCGCCCA-BHQ-1
SOCS2	TCCCTTCCCCGCCAT T	TTGTTTTTGTCGCGGTGATTT	6FAM-CCGAAAAACTCAAAACACCGCAAAATCAT-BHQ1
ACTB	TGGTGATGGAGGAGGTTTAGTAAGT	AACCAATAAAACCTACTCCTCCCTTAA	6FAM-ACCACCACCCAACACACAATAACAAACACA-BHQ-1

Table 3: Methylated genes and high-risk HPV DNA in patients with no dysplasia/low-grade SIL, high-grade SIL and invasive cervical cancer

Variables	No dysplasia/low-grade SIL n=13	High-grade SIL n=31	Invasive Cancer n=5	*p*
Genes				
SOCS1 methylated	0%	7%	60%	< 0.001 (*)
CDH1 methylated	23%	39%	100%	0.011 (*)
TIMP3 methylated	0%	16%	100%	< 0.001
GSTP1 methylated	0%	7%	20%	0.282
DAPK methylated	8%	23%	80%	0.006 (*)
hTERT methylated	0%	0%	80%	< 0.001 (*)
CDH13 methylated	8%	13%	100%	< 0.001 (*)
HSPA2 methylated	0%	3%	60%	< 0.001 (*)
MLH1 methylated	0%	3%	40%	0.004
RASSF1A methylated	8%	0%	40%	0.002
SOCS2 methylated	23%	45%	60%	0.26
At least one gene methylated	39%	58%	100%	0.06
High-risk HPV DNA positive	15%	68%	60%	0.006
At least one gene methylated and/or high-risk HPV DNA-positive	46%	94%	100%	0.001

Table 4: Methylation of *CDH1* and *CDH13* in serum samples of cervical cancer patients

Characteristics	n[a]	CDH1	CDH13
Stage			
FIGO I	23	30%	0%
FIGO II	24	29%	4%
FIGO III	33	67%[b]	6%
FIGO IV	13	23%	8%
Tumor grade			
1	22	50%	0%
2	50	34%	6%
3	16	56%	6%
Histology			
squamous	84	42%	5%
adeno/adenosquamous	9	44%	0%
Age			
< 50	38	36%	5%
≥ 50	55	46%	4%
Total	93	42%	4%

Tumor grade was unknown in five cases

[a] n, number of cases examined.

[b] P = 0.005 (chi^2 test)

Table 5: Multivariate analysis for risk of relapse

Variable	Relative risk of relapse (95% CI)	P value
Stage		**<0.001**
FIGO II (vs. FIGO I)	0.6 (0.2-1.6)	0.316
FIGO III (vs. FIGO I)	1.3 (0.6-2.9)	0.524
FIGO IV (vs. FIGO I)	8.6 (3.3-22.3)	**<0.001**
Tumor grade		0.132
grade 2 (vs. grade 1)	1.9 (0.9-4.2)	0.092
grade 3 (vs. grade 1)	2.1 (0.9-4.8)	0.065
Histology		
squamous (vs. adeno/adenosquamous)	0.6 (0.2-1.7)	0.303
Age	0.97 (0.95-0.99)	**0.012**
CDH1 and/or *CDH13* methylated (vs. un-methylated)	2.5 (1.3-4.6)	**0.005**

Table 6: Multivariate analysis for overall survival

Variable	Relative risk of death (95% CI)	P value
Stage		**<0.001**
FIGO II (vs. FIGO I)	0.7 (0.3-1.8)	0.504
FIGO III (vs. FIGO I)	1.0 (0.4-2.5)	0.993
FIGO IV (vs. FIGO I)	11.1 (4.0-30.4)	**<0.001**
Tumor grade		0.07
grade 2 (vs. grade 1)	2.3 (1.1-5.0)	**0.037**
grade 3 (vs. grade 1)	2.5 (1.0-6.1)	**0.041**
Histology		
squamous (vs. adeno/adenosquamous)	1.3 (0.4-4.3)	0.645
Age	1.0 (0.9-1.0)	0.146
CDH1 and/or *CDH13* methylated (vs. un-methylated)	2.5 (1.3-4.8)	**0.005**

Table 7: Genes and equivalent sequences according to the invention.

Gene/Associated gene	Genomic sequence SEQ ID NO:	Converted sequence SEQ ID NO:
CDH1	1	12, 13, 34 and 35
CDH13	2	14, 15, 36 and 37
RASSF1A	3	16, 17, 38 and 39
hMLH1	4	18, 19, 40 and 41
HSPA2	5	20, 21, 42 and 43
SOCS1	6	22, 23, 44 and 45
SOCS2	7	24, 25, 46 and 47
GSTP1	8	26, 27, 48 and 49
DAPK1	9	28, 29, 50 and 51
TIMP3	10	30, 31, 52 and 53
hTERT	11	32, 33, 54 and 55
SFRP2	64	68, 69, 76 and 77
SFRP4	65	70, 71, 78 and 79
SFRP5	66	72, 73, 80 and 81
CCND2	67	74, 75, 82 and 83

Table 8

HUGO Gene Nomenclature	Location of Amplicon in Gene	Forward Primer Sequence	Reverse Primer Sequence	Probe Oligo Sequence
SFRP2	Promoter	AAACCTACCCGCC CGAAA (SEQ ID NO 84)	GTTGAACGGTGGTT GGAGATTC (SEQ ID NO 85)	CGCCTCGACGAACTTCGT TTTCCCT (SEQ ID NO 86)
SFRP4	Promoter/5'UTR	TCCGCCGTCTAAC ACACAAA (SEQ ID NO 87)	TTCGTAATGGTCGT GGTTGGT (SEQ ID NO 88)	CAACGCCAACTCTCAACCT TCGAAACG (SEQ ID NO 89)
SFRP5	Promoter/5'UTR	GAACGCCCCGACT AATCCTAA (SEQ ID NO 90)	TAGGCGGTCGGAG ATTGGT (SEQ ID NO 91)	CTCCCACCTCGAAACTCCA ACCCG (SEQ ID NO 92)
CCND2	Promoter	GGAGGGTCGGCG AGGAT (SEQ ID NO 93)	TCCTTTCCCCGAAA ACATAAAA (SEQ ID NO 94)	CACGCTCGATCCTTCGCC CG (SEQ ID NO 95)

Table 9

Patient Age	Disease diagnosis	SFRP1 (PMR)	SFRP2 (PMR)	SFRP5 (PMR)	CCND2 (methyl/unmethyl)
22	ovarian cancer	0,00	0,00	0,03	0,00
82	ovarian cancer	0,49	0,34	0,60	0,00
68	cervical cancer	0,00	0,10	0,89	methyl
61	cervical cancer	0,00	0,65	3,47	0,00
58	cervical cancer	1,49	17,60	3,57	0,00
31	cervical cancer	0,06	1,84	23,16	0,00
37	cervical cancer	0,00	0,01	0,04	0,00
46	cervical cancer	258,32	39,58	0,87	methyl
75	endometrial cancer	0,00	28,51	0,76	0,00
60	endometrial cancer	0,09	1,21	1,37	0,00
72	endometrial cancer	0,00	0,15	1,38	methyl
59	endometrial cancer	0,00	16,15	5,35	0,00
17	non-neoplastic	0,00	0,06	0,32	0,00
70	non-neoplastic	0,00	0,00	1,04	0,00
55	non-neoplastic	0,00	0,00	0,00	0,00
56	non-neoplastic	2,05	0,54	1,83	0,00
46	non-neoplastic	0,00	0,00	0,04	0,00
29	non-neoplastic	0,00	0,00	0,00	0,00
63	non-neoplastic	0,14	1,26	1,00	0,00
56	non-neoplastic	0,00	0,00	0,00	0,00
75	non-neoplastic	0,00	0,79	0,53	0,00
69	non-neoplastic	0,00	0,12	0,61	0,00
60	non-neoplastic	0,00	0,00	0,00	0,00
36	non-neoplastic	0,00	0,00	2,96	0,00